



ETIČKI ASPEKTI SISTEMA ZA PREPORUČIVANJE

Marko Vuković

SADRŽAJ

- ❑ Teorijske osnove sistema za preporučivanje
- ❑ Teorijske osnove etike
- ❑ Analiza šest glavnih etičkih problema u sistemima za preporučivanje
- ❑ Odgovoran pristup razvoju sistema za preporučivanje
- ❑ Rezime

TEORIJSKE OSNOVE SISTEMA ZA PREPORUČIVANJE





- ❑ Sistemi za preporučivanje predstavljaju klasu aplikacija čiji je cilj da pruže korisniku preporuku o potencijalno zanimljivom proizvodu, usluzi ili sadržaju.
- ❑ Prema načinu na koji daju preporuke, mogu se podeliti u dve grupe:
 1. Sistemi zasnovani na sadržaju (engl. *content-based recommender systems*)
 2. Sistemi zasnovani na kolaborativnom filtriranju (engl. *collaborative filtering recommender systems*)
- ❑ Sistemi zasnovani na sadržaju daju preporuke na osnovu osobina koje karakterišu korisnike i sadržaj.
- ❑ Sistemi zasnovani na kolaborativnom filtriranju daju preporuke na osnovu sličnosti između korisnika i/ili sadržaja koji se preporučuje.

TEORIJSKE OSNOVE SISTEMA ZA PREPORUČIVANJE

- ❑ U sistemima za preporučivanje postoje dve klase entiteta:
 1. Korisnici
 2. Sadržaj koji se preporučuje
- ❑ Određeni korisnici preferiraju određenu vrstu sadržaja, a glavni zadatak sistema za preporučivanje je da te preferencije detektuje iz podataka.





TEORIJSKE OSNOVE SISTEMA ZA PREPORUČIVANJE

- ❑ U sistemima za preporučivanje, podaci se prezentuju kao matrice ocene korisnika (engl. *utility matrix*).
- ❑ Redovi ove matrice predstavljaju korisnike, dok kolone predstavljaju stavke sadržaja za preporučivanje.
- ❑ Presek reda i kolone predstavlja ocenu koju je korisnik dao određenoj stavki.
- ❑ Matrica je u često u praksi slabo popunjena (engl. *sparse*)

	FILM #1	FILM #2	FILM #3	FILM #4
	2	?	4	5
	5	5	4	3
	1	3	5	2
	4	3	?	5

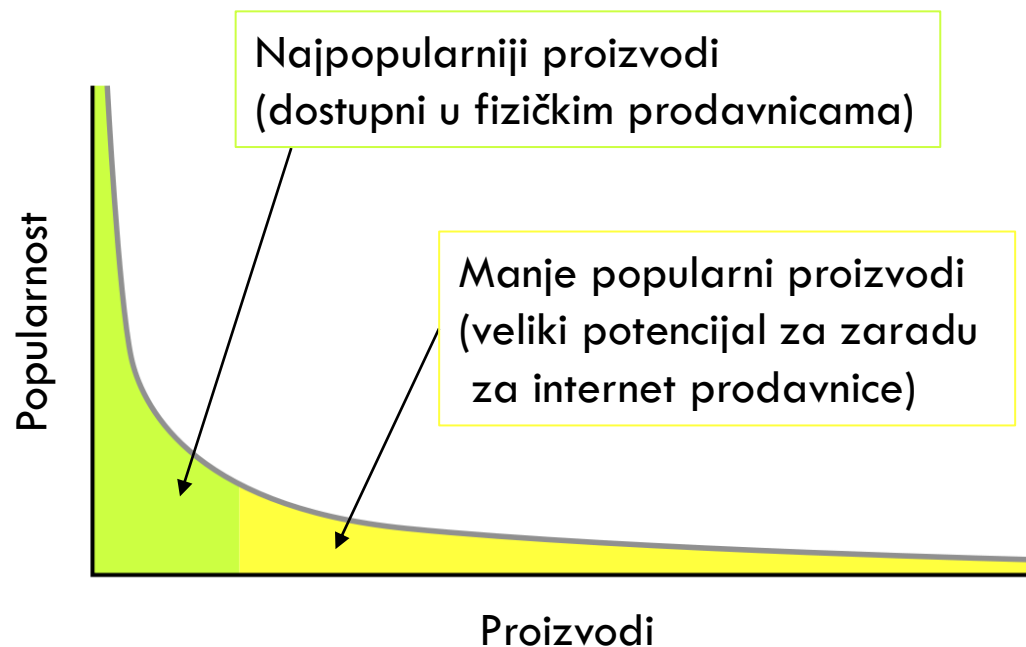
TEORIJSKE OSNOVE SISTEMA ZA PREPORUČIVANJE

- ❑ Cilj sistema za preporučivanje je da predvidi nedostajuće vrednosti matrice ocena korisnika.
- ❑ Nije neophodno predvideti sve vrednosti, već samo one za koje se pretpostavlja da će biti visoke.

	FILM #1	FILM #2	FILM #3	FILM #4
	2	?	4	5
	5	5	4	3
	1	3	5	2
	4	3	?	5

TEORIJSKE OSNOVE SISTEMA ZA PREPORUČIVANJE

- ❑ Nameće se pitanje: zašto su nam uopšte potrebni sistemi za preporučivanje?
- ❑ Odgovor na to pitanje daje fenomen dugog repa.
- ❑ Nepostojanje ograničenja prostora omogućava internet prodavnicama da svojim korisnicima ponude znatno veći broj proizvoda od fizičkih prodavnica.
- ❑ Sistemi za preporučivanje služe za preporučivanje proizvoda potencijalnim kupcima.



TEORIJSKE OSNOVE SISTEMA ZA PREPORUČIVANJE

SISTEMI ZASNOVANI NA SADRŽAJU

- ❑ Preporučivanje u sistemima zasnovanim na sadržaju zasniva se na konstruisanju profila korisnika i stavki sadržaja za preporučivanje.
- ❑ Profili stavki se predstavljaju popoću njihovih važnih karatkeristika, dok se profili korisnika formiraju kao agragacija stavki koje se korisniku dopadaju.
- ❑ Profili se najčešće predstavljaju kao vektori u prostoru važnih karakteristika stavki.
- ❑ Preporučivanje se vrši pronalaženjem onih stavki čiji vektori imaju najmanju udaljenost od vektora korisnika.

TEORIJSKE OSNOVE SISTEMA ZA PREPORUČIVANJE

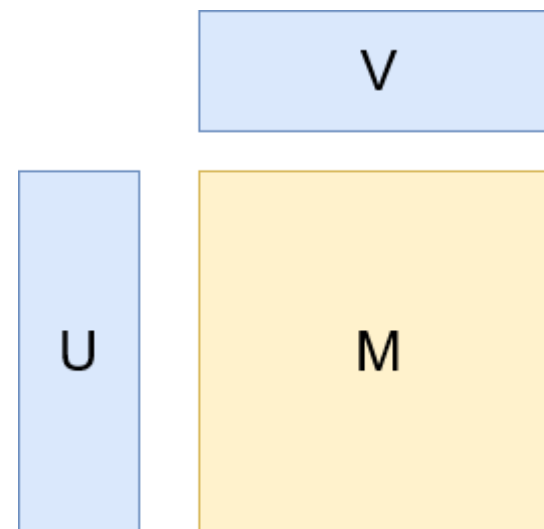
SISTEMI ZASNOVANI KOLABORATIVNOM FILTRIRANJU

- ❑ Preporučivanje u sistemima zasnovanim na kolaborativnom filtriranju zasniva se na kombinovanju ocena većeg broja korisnika, kako bi se odredile nedostajuće vrednosti iz matrice ocena korisnika.
- ❑ Postoje dva glavna tipa pristupa kolaborativnom filtriranju:
 1. Pristupi zasnovani na memoriji
 2. Pristupi zasnovani na modelima
- ❑ Pristupi zasnovani na memoriji koriste ocene iz matrice ocena korisnika kako bi izračunali sličnost između korisnika ili stavki. Postoje dva glavna pristupa:
 1. Pristup pronalaženja sličnih korisnika
 2. Pristup pronalaženja sličnih stavki

TEORIJSKE OSNOVE SISTEMA ZA PREPORUČIVANJE

SISTEMI ZASNOVANI KOLABORATIVNOM FILTRIRANJU

- ❑ Pristupi zasnovani na modelima zasnovani su na upotrebi različitih algoritama mašinskog učenja.
- ❑ Jedan primer je redukcija dimenzionalnosti.
- ❑ Kod redukcije dimenzionalnosti, matrica ocena korisnika M razbija se na dve duge, ali tanke matrice (U i V), čiji proizvod dobro aproksimira polaznu matricu.
- ❑ Ova tehnika naziva se UV-dekompozicija.



TEORIJSKE OSNOVE SISTEMA ZA PREPORUČIVANJE

NETFLIX IZAZOV

- ❑ Kompanija Netflix, 2006. godine organizovala je izazov za poboljšanje sistema za preporučivanje CineMatch.
- ❑ Ponudili su nagradu od 1.000.000 dolara najboljem rešenju, ukoliko to rešenje nadmaši tačnost CineMatch algoritma za 10%.
- ❑ Pobedničko rešenje donelo je tačnost od 10.10%, i kombinovalo je prethodno opisane tehnike.

TEORIJSKE OSNOVE ETIKE

- ❑ Etika ili moralna filozofija je grana filozofije koja obuhvata sistematizaciju, odbranu i preporučivanje koncepata dobrog i lošeg ponašanja.
- ❑ Etiku čine tri oblasti izučavanja:
 1. Metaetika
 2. Primenjena etika
 3. Normativna etika
- ❑ Metaetika izučava poreklo i značenje etičkih principa.
- ❑ Primenjena etika predstavlja praktičnu primenu moralnog rasuđivanja na konkretne etičke probleme.

TEORIJSKE OSNOVE ETIKE

- ❑ Normativna etika nastoji da izgradi sveobuhvatnu etičku teoriju koja na sistematičan način može da odgovori na pojedinačna etička pitanja.
- ❑ Na osnovu toga šta se primarno vrednuje, normativne etičke teorije mogu se podeliti u tri glavne grupe:
 1. Konsekvencijalizam (koji primarno vrednuje posledice)
 2. Deontološka etika (koja primarno vrednuje pravila)
 3. Aretička etika (koja primarno vrednuje vrline)

TEORIJSKE OSNOVE ETIKE

KONSEKVENCIJALIZAM

- ❑ Konsekvencijalizam je gledište koje primarno vrednuje posledice postupaka.
- ❑ Utilitarizam je teorija u okviru konsekvencijalizma koja posledice vrednuje spram toga koliko doprinose korisnosti, odnosno sreći.

Put do pakla
je popločan
dobrim
namerama.

TEORIJSKE OSNOVE ETIKE

DEONTOLOŠKA ETIKA

- ❑ Deontološka etika označava etičku teoriju koja primarno vrednuje pravila postupanja.
- ❑ Pojedinačni postupci se vrednuju derivativno, spram toga jesu li ili nisu u skladu sa pravilima.

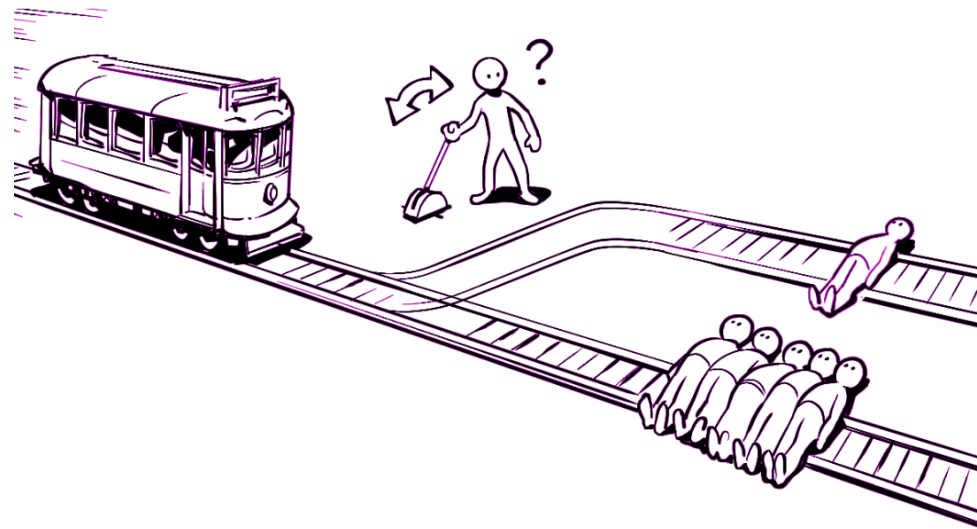
Fiat iustitia,
pereat
mundus.

Neka bude pravda, makar propao svet.

TEORIJSKE OSNOVE ETIKE

MORALNE DILEME – PROBLEM JUREĆEG TRAMVAJA

- ❑ Različite etičke teorije mogu dati različite odgovore na moralne dileme.
- ❑ Konsekvencionalista bi verovatno rekao da bi povukao polugu jer spasavanje pet osoba donosi veću sreću od spasavanja jedne osobe.
- ❑ Sledbenik deontološke etike verovatno ne bi povukao polugu, jer se direktno izazivanje smrti osobe sasvim sigurno kosi sa njegovim moralnim načelima i pravilima.



TEORIJSKE OSNOVE ETIKE

- ❑ Aretička etika ili etika vrlina je oblast etike koja primarno vrednuje karakerne osobine ljudi (vrline i mane), a ne njihove postupke.
- ❑ Konsekvencijalizam i deontološka etika nam govore šta treba da činimo.
- ❑ Aretička etika nam govori kakvi treba da budemo.

ETIČKI PROBLEMI SISTEMA ZA PREPORUČIVANJE

1. PREPORUČIVANJE NEPRIKLADNOG SADRŽAJA

- ☐ Šta zapravo predstavlja neprimeren sadržaj?
 - ☐ Sadržaj neprikladan za određene grupe korisnika (npr. u određenim kulturama)
 - ☐ Sadržaj neprikladan za pojedinačne korisnike
- ☐ Pronađena rešenja ovog problema:
 - ☐ Rešenje problema sadržaja neprikladnog za grupe korisnika upotrebom globalnog etičkog filtera (An Automatic Ethical-Based Recommender System for E-Commerce)
 - ☐ Rešenje problema sadržaja neprikladnog za pojedinačne korisnike upotrebom korisnički podesivog filtera (I should not recommend it to you even if you will like it: the ethics of recommender system)

ETIČKI PROBLEMI SISTEMA ZA PREPORUČIVANJE

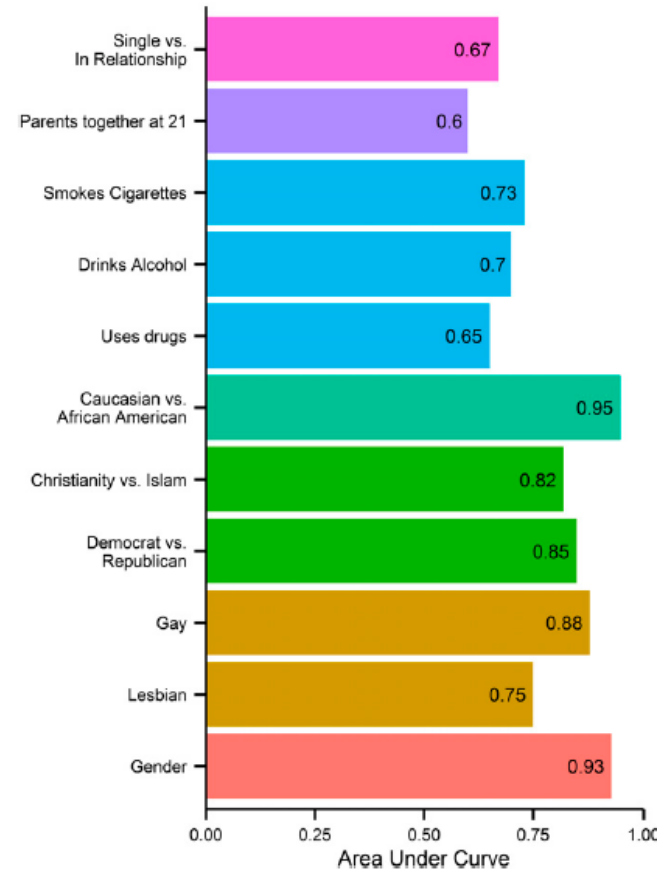
1. PREPORUČIVANJE NEPRIKLADNOG SADRŽAJA

- ☐ Problemi sa upotrebom korisnički podesivog etičkog filtera:
 - ☐ Pitanje cenzure
 - ☐ Implementiranje korisnički podesivog filtera prebacuje odgovornost na korisnike, od kojih neki nemaju ili dovoljno tehničkog znanja ili vremena da ih podese.
 - ☐ Parametri korisnički podesivih filtera predstavljaju metapodatke na osnovu kojih sistem može da donese određene zaključke o korisnicima.

ETIČKI PROBLEMI SISTEMA ZA PREPORUČIVANJE

2. NARUŠAVANJE PRIVATNOSTI KORISNIKA

- Rezultati istrivanja „*Private traits and attributes are predictable from digital records of human behavior*“ pokazali su da se brojne lične informacije koje korisnici nisu podelili sa sistemom mogu odgonetnuti iz naizgled beznačajnih podataka kakvi su Fejsbuk lajkovi.
- Model na osnovu Fejsbuk lajkova korisnika sa veoma velikom preciznošću pogađa pol, seksualnu orijentaciju, političe i religiozne stavove i rasnu pripadnost korisnika, ali i neke osobine ličnosti, doduše sa nešto nižom preciznošću.



ETIČKI PROBLEMI SISTEMA ZA PREPORUČIVANJE

2. NARUŠAVANJE PRIVATNOSTI KORISNIKA

❑ Član 17. Zakona o zaštiti podataka o ličnosti glasi:

Zabranjena je obrada kojom se otkriva rasno ili etničko poreklo, političko mišljenje, versko ili filozofsko uverenje ili članstvo u sindikatu, kao i obrada genetskih podataka, biometrijskih podataka u cilju jedinstvene identifikacije lica, podataka o zdravstvenom stanju ili podataka o seksualnom životu ili seksualnoj orijentaciji fizičkog lica.

Izuzetno, obrada iz stava 1. ovog člana dopuštena je u sledećim slučajevima:

1. lice na koje se podaci odnose je dalo izričit pristanak za obradu za jednu ili više svrha obrade, osim ako je zakonom propisano da se obrada ne vrši na osnovu pristanka;

ETIČKI PROBLEMI SISTEMA ZA PREPORUČIVANJE

2. NARUŠAVANJE PRIVATNOSTI KORISNIKA

- ❑ Na koji način je privatnost izložena riziku u sistemima za preporučivanje?
 1. Kada se podaci prikupljaju ili dele bez pristanka korisnika
 2. Kada se podaci skladište postoji rizik da će biti otkriveni trećim licima ili biti žrtva pokušaja procesa de-anonimizacije (Netflix-ov izazov)
 3. Kada sistem može da (omogući nekom da) donese zaključke o korisniku na osnovu podataka koje poseduje
 4. Rizici se javljaju i zbog same prirode kolaborativnog filtriranja, koja omogućuje sistemu da donosi zaključke i o onim korisnicima o kojima ima manje podataka, zbog interakcije drugih korisnika sa sistemom.

ETIČKI PROBLEMI SISTEMA ZA PREPORUČIVANJE

2. NARUŠAVANJE PRIVATNOSTI KORISNIKA

- ❑ Na koji način je se može smanjiti rizik kojem je privatnost izložena u sistemima za preporučivanje?
 1. Upotrebom arhitektura koje podržavaju privatnost korisnika
 2. Algoritamskim rešenjima
 3. Uvođenjem zakonske regulative i kodeksa profesionalne etike
 4. Implementiranje korisničkih podešavanja za očuvanje privatnosti
- ❑ Korisnički orijentisan pristup se suočava sa istim problemima kao i u prethodnim slučajevima.

ETIČKI PROBLEMI SISTEMA ZA PREPORUČIVANJE

2. NARUŠAVANJE PRIVATNOSTI KORISNIKA

- ❑ Na koji način privatnost može da utiče na sisteme za preporučivanje?
- ❑ Svojevrsnu Scilu i Haribdu sistema za preporučivanje mogu da predstavljaju društvena prihvatljivost i zakonska regulativa.
- ❑ Problem privatnosti u sistemima preporučivanje mora biti rešen upotrebom makroetičkog pristupa (razmatranjem etike podataka, etike algoritama i etike prakse).



ETIČKI PROBLEMI SISTEMA ZA PREPORUČIVANJE

3. UTICAJ NA AUTONOMIJU

- ❑ Uticaj na autonomiju korisnika je u samom prirodi sistema za preporučivanje – preporučivanje ili gurkanje ka nekoj od opcija sistemi za preporučivanje pokušavaju da „navuku“ korisnike na određenu vrstu sadržaja ili im uskraćuju opcije.
- ❑ Ovakve akcije mogu se protumačiti na više različitih načina:
 - ❑ kao benigne,
 - ❑ kao akcije upitne prirode, i
 - ❑ kao maliciozne.

ETIČKI PROBLEMI SISTEMA ZA PREPORUČIVANJE

4. HERMETIČNOST (ZATVORENOST)

- ❑ Hermetičnost sistema za preporučivanje obuhvata niz srodnih problema koji bi mogli biti rešeni obrazlaganjem preporuka koje sistem za preporučivanje daje.
- ❑ Hermetičnost je blisko povezana sa problemima autonomije i uticaja na društvo.
- ❑ Definisano je sedam ciljeva koje može imati implementiranje obrazloženja uz preporuke (Designing and evaluating explanations for recommender systems):
 1. Transparentnost
 2. Interaktivnost
 3. Povećanje stepena poverenja
 4. Povećanje efektivnosti
 5. Ubeđivanje korisnika
 6. Povećanje efikasnosti
 7. Povećanje zadovoljstva korisnika

ETIČKI PROBLEMI SISTEMA ZA PREPORUČIVANJE

5. PRISTRASNOST

- ❑ Algoritamska nepristrasnost predstavlja sprečavanje algoritama da usvoje predrasude prema korisnika iz podataka nad kojima se obučavaju.
- ❑ Većina istraživanja iz oblasti sistema za preporučivanje zasnovanih na kolaborativnom filtriranju, zasnovano je na pretpostavci da nedostajuće ocene u matrici ocena korisnika ne zavise od vrednosti tih ocena ili vrednosti bilo kojih drugih ocena.
- ❑ Ipak, određena istraživanja (Collaborative prediction and ranking with non-random missing data) su pokazala da to u praksi nije baš tako – ljudi su spremniji da ocene onaj sadržaj koji im se dopao, pre nego onaj koji im se nije dopao.

ETIČKI PROBLEMI SISTEMA ZA PREPORUČIVANJE

5. PRISTRASNOST

- ❑ Postoji statistička teorija koju su stvorili Litl i Ruben, koja kaže da neispravne pretpostavke o nedostajućim podacima mogu uvesti pristransost (engl. *bias*) u parametre i predikciju modela.
- ❑ Sistem za preporučivanje izbornih predmeta na jednom evropskom univerzitetu lepo opisuje suštinu problema pristrasnosti u podacima i modelima, koji rezultuju diskriminacijom određenih grupa studenata (Collaborative prediction and ranking with non-random missing data).
- ❑ Zbog čega dolazi do pojave pristrasnosti i favorizacije u skupovima podataka?
 - ❑ Populaciona nebalansiranost
 - ❑ Posmatračka pristrasnost

ETIČKI PROBLEMI SISTEMA ZA PREPORUČIVANJE

5. PRISTRASNOST

□ Pet metrika pristrasnosti pomoću kojih se pristrasnost može izbeći prilikom obučavanja modela:

1. Vrednosna pristrasnost
2. Apsolutna pristrasnost
3. Koeficijent potcenjivanja
4. Koeficijent precenjivanja
5. Mera nejednakosti

ETIČKI PROBLEMI SISTEMA ZA PREPORUČIVANJE

6. UTICAJ NA DRUŠTVO

- ❑ Sistemi za preporučivanje utiču na način na koji korisnici konzumiraju sadržaj i na način na koji formiraju mišljenje o svetu koji ih okružuje.
- ❑ Misli se pre svega na sisteme za preporučivanje novinskih članaka i objava na društvenim mrežama koji formiraju „ideološke balončiće“ oko svojih korisnika, i na taj način sprečavaju javnu diskusiju i utiču na stvaranje i produbljivanje predrasuda.
- ❑ Problem predstavljaju i male grupe aktivnih korisnika koje svojim interakcijama sa sistemima za preporučivanje uspevaju da lažne vesti i teorije zavera dovedu iz zabačenih kutaka interneta do velike medijske pažnje, kao i da ih pretvore u kanale za prenošenje političke propagande.

ETIČKI PROBLEMI SISTEMA ZA PREPORUČIVANJE

6. UTICAJ NA DRUŠTVO

□ Dva pristupa rešavanja problema polarizacije i manipulacije u društvu:

1. U skladu sa interesima korisnika
2. U skladu sa interesima društva

ODGOVORAN PRISTUP RAZVOJU SISTEMA VEŠTAČKE INTELIGENCIJE

- ❑ Pristup je opisan u knjizi Odgovorna veštačka inteligencija Virdžinije Dignum.
- ❑ Opisan je pristup razvoju sistema veštačke inteligencije koji obuhvata inženjerski, filozofski i sociološki pristup, a cilj predstavlja implementiranje sistema kojim će biti zadovoljni i korisnici i ponuđači sadržaja, i sistema koji će poštovati moralna načela svih korisnika.
- ❑ Sistemi za preporučivanje predstavljaju jednu od oblasti veštačke inteligencije, te se ceo pristup može u potpunosti primeniti i na njihov razvoj.
- ❑ Odgovoran pristup je neophodan ne samo kako bi se osiguralo da se sistem razvija na dobar način, već i da se razvija iz dobrih razloga i da njegova upotreba neće doneti negativne posledice po korisnike, društvo i okolinu.

ODGOVORAN PRISTUP RAZVOJU SISTEMA VEŠTAČKE INTELIGENCIJE

- ❑ Odgovoran pristup razvoju sistema veštačke inteligencije zasnovan je na Odgovornom pristupu istraživanju i inovaciji (engl. *Responsible research and innovation – RRI*) – konceptu Evropske unije koji proces istraživanja i inovacije predstavlja kao interaktivan proces u kom predstavnici društva i inovatori međusobno razgovaraju o temama (etičke) prihvatljivosti, održivosti i društvene poželjnosti inovativnih i proizvodnih procesa.
- ❑ Autorka predlaže pristup koji naziva Dužnost, odgovornost i transparentnost (engl. *Accountability, Responsibility and Transparency – ART*)
- ❑ ART pristup zasnovan je na principu dizajniranja za vrednosti.

ODGOVORAN PRISTUP RAZVOJU SISTEMA VEŠTAČKE INTELIGENCIJE

- Dizajniranje za vrednosti pruža podlogu za održavanje sledećih procesa:
 1. Identifikacije zainteresovanih strana
 2. Otkrivanja vrednosti i zahteva zainteresovanih strana
 3. Agregacija i interpretacija vrednosti zainteresovanih strana
 4. Praćenje formalnih veza između vrednosti, interpretacija vrednosti i odabir sistemskih funkcionalnosti
 5. Pružanje podrške za odabir sistemskih komponenti

ODGOVORAN PRISTUP RAZVOJU SISTEMA VEŠTAČKE INTELIGENCIJE

- ❑ Sisteme veštačke inteligencije i sisteme za preporučivanje potrebno je posmatrati kao socio-tehnološke sisteme.
- ❑ Značenje pojmova odgovornost, dužnost i transparentnost iz samog naziva ART pristupa:
 1. Dužnost podrazumeva dužnost sistem da objasni i opravda svoje odluke i akcije
 2. Odgovornost se odnosi na odgovornost koje ljudi zaduženi za razvijanje ovih sistema moraju biti svesni
 3. Transparentnost se odnosi na sposobnost da se opiše, istraže i ponove mehanizmi putem kojih sistemi veštačke inteligencije donose odluke i obučavaju se, ali i na transparentnost porekla podataka, procesa razvoja i zahteva zainteresovanih strana

ODGOVORAN PRISTUP RAZVOJU SISTEMA VEŠTAČKE INTELIGENCIJE

□ Primena pristupa u praksi podrazumeva sledeće:

1. Uređivanje aktivnosti istraživanja, razvoja i stavljanja u upotrebu sistema veštačke inteligencije u društvenom kontekstu
2. Odgovorno ponašanje inženjera i drugih zaposlenih, na individualnom i kolektivnom nivou
3. Rešavanje problema inkluzije, raznovrsnosti i jednakog pristupa
4. Promenu narativa o sistemima veštačke inteligencije

ODGOVORAN PRISTUP RAZVOJU SISTEMA VEŠTAČKE INTELIGENCIJE

1. REGULISANJE PROCESA

- ☐ Donošenje pravnih regulativa o upotrebi veštačke inteligencije u pojedinačnim oblastima.
- ☐ Nezavise i pouzdane institucije bi mogle da validiraju i testiraju algoritme, aplikacije i proizvode i da garantuju za kvalitete sistema izdavanjem sertifikata.

ODGOVORAN PRISTUP RAZVOJU SISTEMA VEŠTAČKE INTELIGENCIJE

2. ODGOVORNO PONAŠANJE

- Garantuje se izradom kodeksa profesionalne etike.
- Kodeksi bi predstavljali javnu izjavu kojom bi se profesionalci i kompanije izjašnjavale o sledećim temama:
 1. O praksi, ophođenju i etičkim načelima svih uključenih u razvoj sistema
 2. O očekivanjima od profesije i zajednice
 3. Davali bi izjavu društvu šta da očekuje
 4. Omogućila bi profesionalcima da u skladu s njom donose dobre odluke, kao i da ih na taj način pravdaju

ODGOVORAN PRISTUP RAZVOJU SISTEMA VEŠTAČKE INTELIGENCIJE

3. INKLUZIJA, RAZNOVRSNOST I JEDNAKOST PRISTUPA

- ❑ Razvojni timovi treba da uključe i sociologe, filozofe i druge.
- ❑ Razvojni timovi treba da poštuju rodne, etničke i kulturne razlike.
- ❑ Potrebno je obogatiti obrazovanje svih uključenih u proces razvoja sistema, kako inženjera, tako i pravnika ali i drugih.

ODGOVORAN PRISTUP RAZVOJU SISTEMA VEŠTAČKE INTELIGENCIJE

3. INKLUZIJA, RAZNOVRSNOST I JEDNAKOST PRISTUPA

□ Obogačenje inženjerskih studijskih programa podrazumeva:

1. Analizu distribuirane prirode veštačke inteligencije i kompleksnosti interakcije ljudi sa sistemima
2. Refleksija o uticaju sistema na društvo
3. Refleksija o tome koliko je teško proceniti uticaj odluka na individualnom i sistemskom nivou na ljudska prava, demokratiju i obrazovanje
4. Posledice nedostatka inkluzije i raznovrsnosti
5. Razumevanje pravnih regulativa
6. Razumevanje društvenih, pravnih i ekonomskih modela socio-tehnoloških sistema
7. Isticanje važnosti vrednosno zasnovanih pristupa razvoju

ODGOVORAN PRISTUP RAZVOJU SISTEMA VEŠTAČKE INTELIGENCIJE

4. PROMENA NARATIVA

- Odgovornost u veštačkoj inteligenciji počinje sa adekvatnim narativom, koji demistifikuje mogućnosti, domete i procese i omogućava svima da učestvuju u razgovoru o ulozi koju veštačka inteligencija ima u društvu.

REZIME

☐ Predstavljene su osnove sistema za preporučivanje:

1. Sistemi zasnovani na sadržaju
2. Sistemi zasnovani na kolaborativnom filtriranju
 1. Sistemi zasnovani na memoriji
 2. Sistemi zasnovani na modelima

☐ Predstavljene su osnove etike:

1. Metaetika
2. Primenjena etika
3. Normativna etika
 1. Konsekvencijalizam
 2. Deontološka etika
 3. Aretička etika

REZIME

□ Razmatrani su etički problemi:

1. Preporučivanje neprikladnog sadržaja
2. Uticaj sistema za preporučivanje na privatnost korisnika
3. Uticaj sistema za preporučivanje na autonomiju korisnika
4. Hermetičnost sistema za preporučivanje
5. Pristrasnost sistema za preporučivanje
6. Uticaj sistema za preporučivanje na društvo

REZIME

□ Na samom kraju, predstavljen je jedan odgovoran pristup razvoju sistema veštačke inteligencije, a samim tim i sistema za preporučivanje – ART pristup koji podrazumeva:

1. Pravno i profesionalno regulisanje procesa razvoja sistema
2. Odgovorno ponašanje ljudi i organizacija u procesu razvoja sistema
3. Rešavanje problema inkluzije, raznovrsnosti i jednakog pristupa
4. Promena narativa o sistemima

LITERATURA

1. Milano, Silvia, Mariarosaria Taddeo, and Luciano Floridi. "Recommender systems and their ethical challenges." *Ai & Society* 35.4 (2020): 957-967.
2. Souali, Kamal, Abdellatif El Afia, and Rdouan Faizi. "An automatic ethical-based recommender system for e-commerce." *2011 International Conference on Multimedia Computing and Systems*. IEEE, 2011.
3. Tang, Tiffany Ya, and Pinata Winoto. "I should not recommend it to you even if you will like it: the ethics of recommender systems." *New Review of Hypermedia and Multimedia* 22.1-2 (2016): 111-138.
4. Kosinski, Michal, David Stillwell, and Thore Graepel. "Private traits and attributes are predictable from digital records of human behavior." *Proceedings of the national academy of sciences* 110.15 (2013): 5802-5805.

LITERATURA

5. Floridi, Luciano and Taddeo, Mariarosaria. „What is data ethics?." *Philosophical Transactions of The Royal Society A Mathematical Physical and Engineering Sciences*. 374. (2016): 20160360. 10.1098/rsta.2016.0360.
6. Friedman, Arik, et al. "Privacy aspects of recommender systems." *Recommender systems handbook*. Springer, Boston, MA, 2015. 649-688.
7. De Vries, Katja. "Identity, profiling algorithms and a world of ambient intelligence." *Ethics and information technology* 12.1 (2010): 71-85.
8. Tintarev, Nava, and Judith Masthoff. "Designing and evaluating explanations for recommender systems." *Recommender systems handbook*. Springer, Boston, MA, 2011. 479-510.

LITERATURA

9. Burke, Robin. "Multisided fairness for recommendation." *arXiv preprint arXiv:1707.00093* (2017).
10. Marlin, Benjamin M., and Richard S. Zemel. "Collaborative prediction and ranking with non-random missing data." *Proceedings of the third ACM conference on Recommender systems*. 2009.
11. Yao, Sirui, and Bert Huang. Collaborative prediction and ranking with non-random missing data." *Advances in neural information processing systems* 30 (2017).
12. Bozdag, Engin, and Jeroen Van Den Hoven. "Breaking the filter bubble: democracy and design." *Ethics and information technology* 17.4 (2015): 249-265.

LITERATURA

14. Dignum, Virginia. *Responsible artificial intelligence: how to develop and use AI in a responsible way*. Springer Nature, 2019.
15. Berčić, Boran. *Filozofija – Sažeto E-izdanje*. Ibis grafika – Zagreb, 2012.
16. Mining of massive datasets class lectures and materials – Stanford University

IZVORI SLIKA, ILUSTRACIJA I GRAFIKA

1. https://www.freepik.com/free-vector/hand-drawn-collection-different-profile-icons_17786264.htm
2. https://www.wikiwand.com/en/Long_tail
3. <https://naukakrozprice.rs/problem-tramvaja/>
4. <https://www.pnas.org/doi/pdf/10.1073/pnas.1218772110>
5. https://en.wikipedia.org/wiki/Between_Scylla_and_Charybdis